NMR and Molecular Modeling Studies of the Interaction between Wheat Germ Agglutinin and the β-D-Glc*p*NAc-(1→6)-α-D-Man*p* Epitope Present in Glycoproteins of Tumor Cells†

Kristina Lycknert,‡ Malin Edblad,‡ Anne Imberty,§ and Göran Widmalm*,‡

Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden, and CERMAV-CNRS (affiliated with Université Joseph Fourier), BP53, F-38041 Grenoble Cedex 09, France

Received January 13, 2004; Revised Manuscript Received May 21, 2004

ABSTRACT: The β-D-Glc*p*NAc-(1→6)-α-D-Man*p* disaccharide is a constituent of highly branched cell-surface glycoconjugates that are malignancy markers. The conformational preference of the disaccharide β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe in solution has been studied by molecular modeling and NMR spectroscopy including 1D ^{1}H,^{1}H T-ROESY experiments and analysis of $^3J_{H,H}$ of the hydroxymethyl group being part of the glycosidic linkage of the disaccharide, which revealed the relative populations of the ω torsion angle as *gt* = 0.60, *gg* = 0.35, and *tg* = 0.05. Good agreement was obtained between the effective proton−proton distances from the experiment and those obtained by molecular modeling when the flexibility at the ω torsion angle was taken into account. Molecular modeling of the disaccharide in the binding sites of the lectin wheat germ agglutinin indicates that several conformations could be adopted in the bound state. ^{1}H NMR and transfer NOESY experiments confirmed that binding took place, and trans-glycosidic proton−proton interactions indicated that a conformational preference was present in the bound state, as observed by the relative change of the NOEs from H1′ to $H6_{pro-R}$ and $H6_{pro-S}$. STD NMR experiments showed that binding occurred in the region of the *N*-acetyl group of the terminal sugar residue. In addition, the *O*-methyl group received saturation transfer because of the proximity to the protein. ^{1}H,^{1}H NOEs indicated that the two methyl groups were close in space, as observed in only one of the predicted bound conformations. Experimental and theoretical data therefore agree that one conformation with a *gt* conformation of the hydroxymethyl group and a negative sign for the ψ torsion angle is indeed selected by the lectin upon binding.

Complex oligosaccharides present in *N*-glycoproteins portray structural variations, also called glycoforms, with some of them being characteristic of surfaces of metastatic cancer cells (*1*). Among these malignancy markers, a large interest has been devoted to the appearance of additional branches, resulting in tri- and tetra-antennary structures (*2*). In such cases, the trimannosyl core of the Asn-linked oligosaccharide contains a β-D-Glc*p*NAc-(1→6)-α-D-Man*p* saccharide. The correlation between neoplastic and malignant transformation and the increase in β-(1→6) branching on Asn-linked carbohydrates is a common observation (*3*, *4*). The presence of this carbohydrate motif can be used as a marker of a malignant phenotype in human breast and colon cancer (*5*, *6*) and in gliomas (*7*). This structural modification is recognizable by a leukoagglutinating lectin from *Phaseolus vulgaris* (L-PHA), which binds with high affinity to galactose-containing β-(1→6)-antennae (*8*).

Wheat germ agglutinin (WGA)[1] is another lectin that has been found to agglutinate malignant cells specifically (*9*). WGA consists of three isolectins, namely, WGA1 (35%), WGA2 (55%), and WGA3 (10%), with small variations in peptide sequence (5−8 residues in 171 amino acids) (*10*). From the structural study (*11*), monomers consisting of four repeat domains dimerize and therefore define eight independent saccharide-binding sites located at the interface between contacting domains. Because of the 2-fold symmetries, the binding sites present only two architectures, one described as the primary binding site, which is generally fully occupied by carbohydrates in the different crystal structures (*12*−*15*), and the other described as the secondary binding site, that is not or only partially occupied and may have too weak of an affinity for sugars for playing a role in solution (*16*).

There are still uncertainties about the structure of the oligosaccharide epitope that is recognized by WGA on cancer cells. WGA is specific for both terminal *N*-acetyl-D-neuraminic acid (NeuAc) and *N*-acetyl-D-glucosamine (GlcNAc), and the binding parameters have been characterized by a variety of methods (*16*−*19*). The GlcNAc-containing cancer marker that could be recognized has been

† Research was supported by the Swedish Research Council and presented in part at the 19th International Carbohydrate Symposium, San Diego, CA, August 9-14, 1998 (abstract AP 085).

* To whom correspondence should be addressed: Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden. Phone: +46 8 16 37 42. Fax: +46 8 15 49 08. E-mail: gw@organ.su.se.

‡ Stockholm University.

§ CERMAV-CNRS.

[1] Abbreviations: STD, saturation transfer difference; WGA, wheat germ agglutinin.

10.1021/bi0499011 CCC: $27.50
Published on Web 07/07/2004

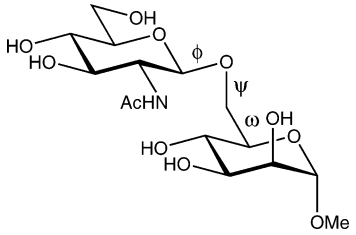

FIGURE 1: Schematic of β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe. Torsion angles relevant for the conformational analysis are indicated as ϕ, ψ, and ω.

proposed to be GlcNAcβ1-6Gal present in poly-*N*-acetyllactosaminoglycans (*20*), and the crystal structure of the complex between WGA-3 and this epitope has been solved (*15*). The GlcNAcβ1-6GalNAc sequence found in core 2 structure *O*-glycans (*21*) could also be a ligand for WGA. Herein, we investigate how WGA binds to the GlcNAcβ1-6Man epitope characteristic of highly branched *N*-glycans, as discussed above. A combination of NMR and molecular modeling approaches are used to delineate the interaction between the lectin and this ligand in the form of an *O*-methyl glycoside, viz., β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe (Figure 1).

MATERIALS AND METHODS

General Procedures. β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe (**1**) was obtained from Carbohydrates International AB, Arlöv, Sweden and WGA of *Triticum vulgaris*, from Sigma Chemical Co., St. Louis, MO. Both materials were used as delivered. Atoms in the β-D-Glc*p*NAc residue are primed, and those in the α-D-Man*p*-OMe residue are unprimed. The torsion angles at the glycosidic linkage between the two sugar residues are denoted by ϕ = O5′−C1′−O6−C6, ψ = C1′−O6−C6−C5, and ω = O6−C6−C5−O5.

NMR Sample Preparation. Disaccharide **1** (5 mg) was lyophilized and subsequently dissolved in D_2O (0.7 mL) to give a concentration of 18 mM. The freeze-dried lectin (3.9 mg) was dissolved in D_2O (2 mL) buffered with sodium phosphate (100 mM at pD 7.2), which resulted in a WGA concentration of 54 μM. The lectin solution (0.7 mL) was added to the disaccharide (1.3 mg), yielding a 4.7 mM solution with respect to **1**. The resulting protein−ligand molar ratio was 1:86, corresponding to ∼1:22 with respect to the primary binding sites of WGA.

NMR Experiments. The NMR experiments were carried out at 300 K on a Bruker DRX 500 MHz spectrometer equipped with a 5-mm PFG triple-resonance CryoProbe and at 298 K on a Varian Inova 600 MHz spectrometer equipped with a 5-mm PFG triple-resonance probe. The ^{1}H and ^{13}C signals were assigned to the corresponding proton and carbon atoms in the disaccharide using 2D ^{1}H,^{1}H DQF−COSY, ^{1}H,^{1}H TOCSY (90-ms mixing time) and gradient-selected ^{1}H,^{13}C HSQC experiments at the 600 MHz spectrometer. Chemical shifts are reported in parts per million relative to internal sodium 4,4-dimethyl-4-sila(2,2,3,3-D_4)pentanoate (δ_H = 0.00) or external dioxane in D_2O (δ_C = 67.4).

Proton−proton cross-relaxation rates in the disaccharide were extracted from a series of eight 1D ^{1}H,^{1}H DPFGSE T-ROESY experiments (*22*) at 600 MHz with increasing mixing times (50−400 ms). Selective excitations were enabled using a 40-Hz broad i-Snob-2 shaped pulse (*23*) of 42.5-ms duration. The gradient duration in the initial DPFGSE part was 1 ms, and the strengths were 0.4 and 1.2 G cm^{-1}, respectively. The DPFGSE part of the pulse sequence was followed by a T-ROESY spin lock with $\gamma B_1/2\pi$ = 2.2 kHz. Spectra were accumulated with a spectral width of 3364 Hz using 8192 complex points, sampling 512 transients at each mixing time. The relaxation delay between the transients was 12.2 s, which corresponds to >5 T_1. Prior to Fourier transformation, the FIDs were zero-filled eight times and multiplied with a 1-Hz exponential line-broadening factor. All spectra were baseline-corrected using a first-order correction and integrated using the same integration limits. Normalized integrals at different mixing times were used to obtain ^{1}H,^{1}H cross-relaxation build-up curves. The cross-relaxation rates were extracted by fitting of the build-up curves to second-order polynomials. When possible, each distance reported in Table 1 is an average of the two cross-relaxation rates obtained from excitations at the different proton resonance frequencies of a proton pair, e.g., from H1′ to $H6_{pro-S}$ and vice versa.

Table 1: ^{1}H and ^{13}C Chemical Shifts (Parts per Million) of β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe

residue	1	2	3	4	5	6	Me
β-D-Glc*p*NAc[a]	4.57	3.744	3.57	3.46	3.48	3.766, 3.95	2.04
	102.6	56.4	74.7	70.8	76.8	61.6	23.1
α-D-Man*p*-OMe	4.74	3.93	3.750	3.61	3.721	3.758,[b] 4.20[c]	3.39
	101.8	70.7	71.4	67.6	72.2	70.0	55.5

[a] δ_{CO} = 175.5. [b] $H6_{pro-R}$. [c] $H6_{pro-S}$.

^{1}H,^{1}H NOESY and ^{1}H,^{1}H ROESY spectra were acquired for the two samples at 600 MHz using a mixing time of 300 ms and a sweep width of 2100 Hz in both dimensions. Data were collected with 80 transients using 4096 complex data points and 128 t_1 increments. The TPPI-States procedure was used for frequency discrimination in the indirect dimension (*24*). Zero-filling to 1024 × 8192 complex points in t_1 and t_2, respectively, was applied, and the FID was treated with a 90° shifted sine-bell window function in both dimensions prior to Fourier transformation. In addition, 1D ^{1}H,^{1}H DPFGSE NOESY experiments (*25*) were performed on the lectin sample essentially as described above with 16 384 transients and a mixing time of 300 ms.

The STD spectrum was recorded at 500 MHz with 32 dummy scans, 8192 transients, and 65 k data points over a spectral width of 5000 Hz. The initial part of the experiment consisted of a short relaxation delay of 100 ms. Selective presaturation of the protein was achieved by Gaussian-shaped pulses of 50-ms duration, an irradiation power $\gamma B_1/2\pi$ = 86 Hz, a delay between pulses of 1 ms, and a total saturation time of 5 s. The on-resonance irradiation of the protein was set at a chemical shift of 6.8 ppm, and the off-resonance irradiation was applied at 30 ppm. Subtraction of on- and off-resonance spectra was achieved via phase cycling. A 30-ms spin-lock pulse with a strength of $\gamma B_1/2\pi$ = 5 kHz was applied to remove background protein signals. The HDO signal was suppressed via a WATERGATE scheme using a 3−9−19 pulse sandwich, bracketed by two pulsed field gradients of equal intensity and sign.

Molecular Modeling. The crystal structure of the complex between WGA-1 and chitobiose (*11*) was used as the

starting model. The crystal structure contains amino acids 29−171, and five amino acids were modified (T56P, Q59H, Y66H, A93S, and G171A) to obtain a model of WGA-2. Hydrogen atoms were added on all protein atoms together with partial charges. A GlcNAc residue was taken from the database of 3D structures of monosaccharides (http://www.cermav.cnrs.fr/cgi-bin/monos/monos.cgi) and was inserted in both the primary and the secondary binding sites in the same orientation as the one observed in the crystal (*11*). The atom types and partial charges of the monosaccharides have been specifically developed for protein−carbohydrate interactions (*26*), whereas Pullman partial charges were used for the protein moiety. Energy minimization was performed using the Tripos force field (*27*), with geometry optimization of the sugar and the side chains of amino acids in the binding sites. A distance-dependent dielectric constant was used in the calculations. Energy minimizations were carried out using the Powell procedure until a gradient deviation of 0.05 kcal mol^{-1} $Å^{-1}$ was attained. Resulting hydrogen-bond networks and hydrophobic interactions were checked and found to be in agreement with the reported ones (*28*).

In a second step, an α-Man-OMe residue was added at the anomeric position of GlcNAc, generating a model of the *β*-Glc*p*NAc-(1→6)-α-Man*p*-OMe disaccharide. Conformations around the ϕ and ψ torsion angles were systematically studied with the three low-energy orientations of the third torsion angle at the glycosidic linkage, ω: *gt* = 60°, *gg* = −60°, and *tg* = 180° (*29*). The systematic search of conformational space was performed with the Tripos force field using 5° steps. Energy maps were then drawn as a function of ϕ and ψ. All low-energy conformations of interest (six for each binding site) were then optimized, taking into account partial charges and flexibility of protein side chains. The potential energy maps of the disaccharide per se were also calculated for the same orientation of the ω angle to allow for a comparison to the carbohydrate−protein complex.

RESULTS AND DISCUSSION

Solution Conformation of the β-D-GlcpNAc-(1→6)-α-D-Manp-OMe Disaccharide. Prior to the interaction studies between *β*-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe and WGA, we assess the overall solution conformation of the disaccharide. The available theoretical conformational space has previously been determined. The energy maps were calculated as a function of the glycosidic angles ϕ and ψ (*30*) with the carbohydrate-specific force-field PFOS (*31*). Because hydration is not taken into account in these calculations, the aim was not to predict in detail the solution behavior but to explore the allowed conformational space. The flexibility is large around the ω torsion angle, and the three maps revealed a plateau centered at $\phi \approx -60°$, as anticipated from the exo-anomeric effect, with more flexibility for the torsion angle ψ. Nevertheless, this torsion angle displays an energy preference for the extended antiperiplanar conformation ($\psi = 180°$).

The assignments of ^{1}H and ^{13}C NMR resonances for **1** are given in Table 1. The ^{1}H chemical-shift assignments of $H6_{pro-R}$ and $H6_{pro-S}$ in the mannosyl residue were done by comparison to structurally related (1 → 6)-linked disaccharides (*32*). The proton−proton scalar coupling constants for the 6-substituted hydroxymethyl group were determined by a total line shape analysis (*33*, *34*) of the ^{1}H NMR spectrum, which resulted in $J_{H5,H6pro-R} = 6.46$ Hz, $J_{H5,H6pro-S} = 1.89$ Hz, and $J_{H6pro-R,H6pro-S} = -11.22$ Hz, assuming a negative sign for the latter coupling constant. From the limiting values of $^3J_{H5,H6}$, it is possible to obtain the populations of the staggered conformers of the hydroxymethyl group (*35*). This gives, for ω, *gt* = 0.60, *gg* = 0.35, and *tg* = 0.05. Thus, the major conformational state has the *gt* conformation, and the *tg* conformer is hardly populated.

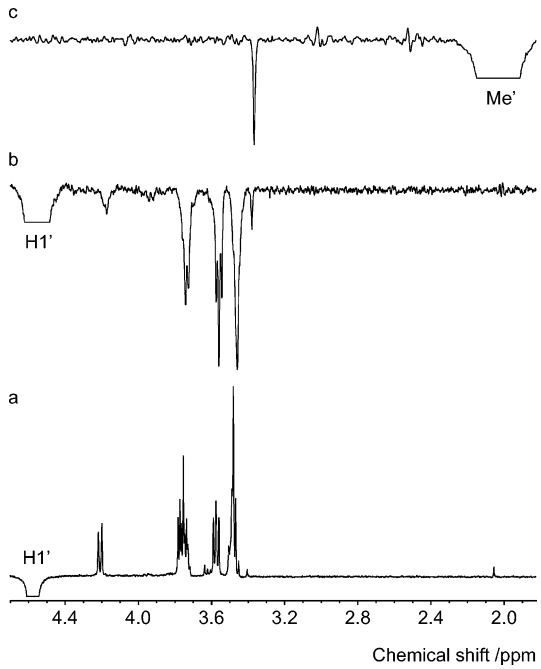

FIGURE 2: (a) One-dimensional ^{1}H,^{1}H T-ROESY spectrum of **1** with selective excitation of the resonance from H1′. One-dimensional ^{1}H,^{1}H NOESY spectrum of **1** in the presence of WGA with selective excitation of the resonance from (b) H1′ and (c) Me′. The mixing time was 300 ms in each of the spectra.

Besides *J*-coupling constants, the NOE is often used in conformational analysis. The ^{1}H,^{1}H T-ROESY experiment is a good alternative to the classical ^{1}H,^{1}H NOESY experiment for small molecules, because of the problem of zero crossing with respect to the ^{1}H,^{1}H NOE per se and signals occurring from TOCSY transfer in the ^{1}H,^{1}H ROESY experiment. One-dimensional analogues of the above experiments are efficient when it is possible to selectively excite a certain resonance. This was employed for the H1′ resonance in **1** in a 1D ^{1}H,^{1}H T-ROESY experiment, and Overhauser effects were observed in the corresponding spectrum (Figure 2a) to, inter alia, H3′, H5′ (used as a reference distance, vide infra), $H6_{pro-R}$, $H6_{pro-S}$, H5, and H4. This also gives an indication of the relative T-ROEs to the H6 protons. The $H6_{pro-S}$ resonance was also selectively excited in another experiment. One-dimensional ^{1}H,^{1}H T-ROESY build-up curves were generated from a series of different mixing times (Figure 3). The initial slope, i.e., the cross-relaxation rate, σ, is proportional to the effective distance between the respective proton pairs. The combined results from different excitations are compiled in Table 2. To obtain an unknown proton−proton distance between the spin pair *i* and *j*, the H1′−H5′ distance was used as a reference and was set to 2.42 Å, obtained from molecular modeling. The isolated spin-pair approximation (ISPA) then makes it possible to extract

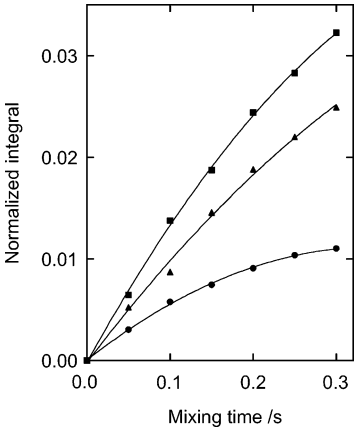

FIGURE 3: ^{1}H,^{1}H cross-relaxation build-up curves obtained for **1** from 1D ^{1}H,^{1}H T-ROESY spectra. H1′−H5′ (■), H1′−H6$_{pro-R}$ (▲), and H1′−H6$_{pro-S}$ (●).

Table 2: Cross-Relaxation Rates from 1D ^{1}H,^{1}H T-ROESY NMR Experiments for Disaccharide **1** and Proton−Proton Distances Derived Using ISPA and Molecular Modeling (see the text)

proton pair	σ (s^{-1})	r_{exp} (Å)	r_{eff} (Å)
H1′−H5′	0.148	2.42[a]	2.42
H1′−H3′	0.091	2.62	2.50
H1′−H6$_{pro-R}$	0.108	2.55	2.33
H1′−H6$_{pro-S}$	0.062	2.80	2.88
H6$_{pro-S}$−H5	0.197	2.31	2.48
H6$_{pro-S}$−H4	0.044	2.96	3.07

[a] Reference distance using a PARM22 force field in CHARMM.

distances by comparing cross-relaxation rates according to (*36*)

$$r_{ij} = r_{ref}(\sigma_{ref}/\sigma_{ij})^{1/6} \quad (1)$$

These results are also shown in Table 2.

To see if the above approximations for the conformational description of the disaccharide are reasonable, we calculate effective proton−proton distances using a molecular mechanics model with ϕ in the exo-anomeric conformation, ψ in the extended antiperiplanar conformation, and ω populating the three staggered conformers. The effective distance r_{eff} is then calculated as

$$1/r_{eff} = (P_{gt}r_{gt}^{-6} + P_{gg}r_{gg}^{-6} + P_{tg}r_{tg}^{-6})^{1/6} \quad (2)$$

in which P is the population of a given conformer and r is the proton−proton distance in the corresponding conformer. Investigation of the H1′−H6$_{pro-S}$, H1′−H6$_{pro-R}$, H6$_{pro-S}$−H5, and H6$_{pro-S}$−H4 pairs, according to eq 2, all showed good agreement to the experimental data (Table 2), lending credence to the analysis performed. In solution, the disaccharide β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe could then be described as a mixture of two populations centered around the two major conformations of ω, namely, *gt* and *gg*.

Binding of WGA to β-D-GlcpNAc-(1→6)-α-D-Manp-OMe. WGA has previously been demonstrated by NMR methods to bind chitooligosaccharides and sialooligosaccharides (*17*, *37*, *38*). In the present paper, the binding of the β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe disaccharide to the protein is demonstrated by the line broadening of the sugar resonances. In particular, an intensity change is present for the methyl group of the *N*-acetyl group (Figure 4a).

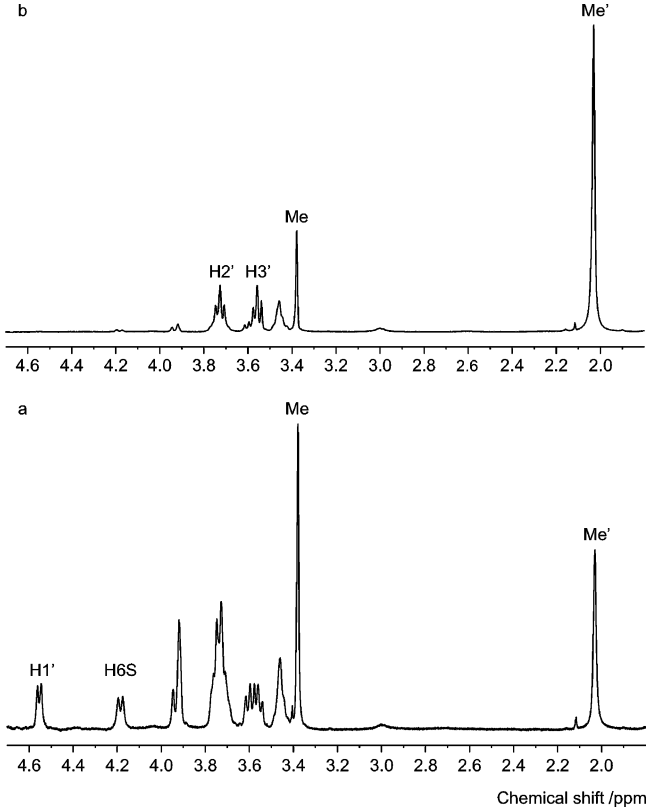

FIGURE 4: (a) ^{1}H NMR spectrum of **1** in the presence of WGA. (b) STD spectrum of **1** in the presence of WGA.

Modeling of the Interaction between β-D-GlcpNAc-(1→6)-α-D-Manp-OMe and WGA. A conformational ϕ/ψ grid search of **1** was performed with the terminal GlcNAc fixed in the primary WGA-binding site (close to Tyr64) and in the secondary WGA-binding site (close to Trp150 of the adjacent monomer). Ramachandran maps were produced that describe the available conformational space (Figure 5) and were compared with the maps of the disaccharide alone that are very similar to the previously calculated ones (*30*). For the three staggered orientations of the ω torsion angle, only a limited subset of the potential energy surface for the free disaccharide is available, indicating that the amino acids of the binding site drastically reduce the flexibility of the β-(1→6) glycosidic linkage of the ligand. In all cases, the main low-energy region of the disaccharide centered around the extended conformation with $\psi \sim 180°$ cannot be fit in the binding site. Only two restricted conformational areas are accessible for the ligand, centered around ψ values of approximately −100° and +100°. Interestingly, very similar results are obtained for the disaccharide in both the primary and secondary binding sites (energy maps not shown).

The conformational ϕ/ψ grid search of the disaccharide in the binding sites therefore yielded six different conformations, corresponding to three values of ω (the *gt*, *gg*, and *tg* conformers) and two values of ψ (approximately −100° and +100°). The final results, after energy minimization, consist of six different docking modes, with different orientations of the mannose residues and therefore different patterns of hydrogen bonds and hydrophobic contacts in these regions. On the other hand, the GlcNAc orientation did not diverge noticeably during optimization and remains similar to the one observed in crystal structures of WGA interacting with GlcNAc (*28*). The energy differences between the binding

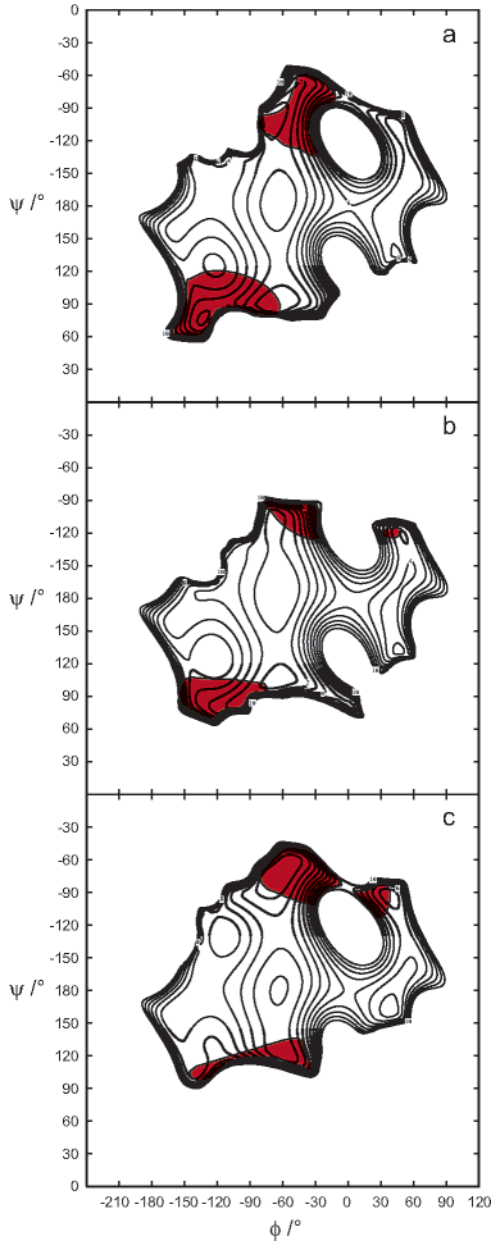

FIGURE 5: Ramachandran maps of **1** for the conformational states of ω: (a) *gt*, (b) *gg*, and (c) *tg*. Accessible conformational regions for **1** in the presence of WGA are shaded in red. The shaded areas represent the contouring of the low-energy conformations in a window up to 20 kcal/mol above the minimum. Typically, the conformations out of the shaded areas are forbidden as a result of steric conflict.

modes are only a few kcal/mol and do not allow us to predict which one corresponds to reality. The characteristics of the models are listed in Table 3, summarizing the different modes of interaction for disaccharide **1** in the binding sites of WGA.

NMR Analysis of the Conformation of β-D-GlcpNAc-(1→6)-α-D-Manp-OMe in the Binding Sites of WGA. ^{1}H,^{1}H NOESY experiments make it possible to investigate the binding of ligands to proteins. When a rapid exchange is present, one obtains negative NOEs from the ligand if it is transiently bound to the protein, thereby reporting on the conformation in the bound state. If any indirect relaxation pathways mediated via the protein are present, they can be identified by aid of a ^{1}H,^{1}H ROESY experiment (*39*). ^{1}H,^{1}H NOESY experiments showed that **1** is binding to WGA. The inherent flexibility at the glycosidic linkage of the (1 → 6)-linked disaccharide, in analogy with other oligosaccharides of a similar kind (*40*, *41*), and the shape of the binding site(s) of WGA may allow for more than one conformation to be recognized. A similar reasoning is possible for the *Aleuria aurantia* agglutinin, a fucosyl-binding protein, when it binds the disaccharide α-L-Fuc*p*-(1→6)-*β*-D-Glc*p*NAc-OMe (*40*). Analysis of trans-glycosidic proton−proton interactions indicate that a conformational preference in **1** is present as observed by the relative change of the NOEs from H1′ to $H6_{pro-R}$ and $H6_{pro-S}$ in a 1D ^{1}H,^{1}H NOESY spectrum (Figure 2b). A ^{1}H,^{1}H ROESY experiment revealed that the cross peaks used in the analysis stemmed from direct interactions within the disaccharide. Furthermore, the presence of cross peaks in the ^{1}H,^{1}H NOESY spectrum between the two methyl groups in **1** as well as in a 1D ^{1}H,^{1}H NOESY spectrum using selective excitation of a methyl resonance (Figure 2c) reveal that the methyl groups should be in close proximity in the bound conformation. This was corroborated by selective excitation of the *O*-methyl resonance resulting in NOEs to Me′ as well as to H1 and H2.

The STD NMR technique is based on the transfer of saturation (*42*) from the protein to the bound ligand (*43*−*45*), which in this case is the disaccharide. Thereby, mapping of the binding epitope can be carried out. The ^{1}H STD NMR spectrum of the mixture (Figure 4b) reveals, as anticipated, a signal from the methyl group of the *N*-acetyl group and resonances from H2′ and H3′ and to some extent H4′/H5′ (overlapped). This result agrees well with the general idea on how WGA recognizes terminal *N*-acetylglucosamine residues (*28*). In addition, saturation transfer also occurs to the *O*-methyl group of the mannosyl residue.

Agreement between NMR and Modeling Work. The modeling approach allows one to define the conformational space available for the disaccharide when bound in the WGA-binding sites. Nevertheless, this prediction needs to be confronted with the NMR results to define, within the possible conformations, the one that is preferred for binding in solution. Analysis of the six conformations modeled in the WGA-binding sites demonstrates that only one of them present a close contact between the two methyl groups of the disaccharide. This corresponds to the model referred to as GT_C1, where ω has the *gt* conformation and $\psi \approx -100°$ (Figure 6). Moreover, in this conformation the $H1'-H6_{pro-R}$ distance is shorter than $H1'-H6_{pro-S}$ (cf. Figure 2b). This particular conformation is from molecular modeling, indeed stabilized by van der Waals contacts between a hydrophobic patch on the mannosyl residue (the H5 and H6 protons) with the aromatic ring of Tyr64, although this is not evident from the STD spectrum.

A comparison can be made with the only structural study of a (1 → 6)-linked disaccharide interaction with WGA (*15*). It is interesting to note that, in three of the four *β*-D-Glc*p*NAc-(1→6)-α-D-Gal*p*-containing ligands bound to WGA-3 as analyzed from crystal data, conformations that can be regarded to belong to a family similar to the conformation found herein were present.

Concluding Remarks. The molecular modeling and NMR approaches agree that the *β*-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe disaccharide displays important flexibility in solution, both around the ω and the ψ torsion angles. As frequently observed in (1 → 6)-linked disaccharides, two conformers

Table 3: Molecular Models Generated by a Conformational Grid Search of β-D-Glc*p*NAc-(1→6)-α-D-Man*p*-OMe in the Presence of WGA Followed by Geometry Optimization[a]

					hydrogen bond		proximity	
model	site	ϕ	ψ	ω	donor	acceptor	proton	AA
GT_C1	P	−67.2	−102.1	57.9			H5, H6	Y64
GT_C2	P	−58.5	87.6	58.6			H1, H4, H6	Y64
GG_C1	P	−62.7	−100.8	−65.0	ManO3	E150b.OE		
GG_C2	P	−86.5	79.4	−64.2	ManO2	GlcNAcO5	H6	Y64
TG_C1	P	−64.0	−75.3	175.4	ManO4	E115b.OE	H4, H6	Y64
TG_C2	P	−56.7	121.8	−178.3	ManO4	H66.NE2	H3, H5	Y64
GT_C1	S	−58.9	−131.2	59.5			H5, H6	W150b
GT_C2	S	−66.8	84.2	57.7			H1, H4, H6	W150b
GG_C1	S	−68.0	−100.5	−64.8	ManO2	D29.OD		
					ManO3	D29.OD		
					W150b.NE	ManO5		
GG_C2	S	−90.3	77.1	−62.3			H6	W150b
TG_C1	S	−62.1	−69.8	176.3	ManO4	D29.OD	H4, H6	W150b
TG_C2	S	−76.7	−152.8	169.7			H3, H5	W150b

[a] Only hydrogen bonds and the proximity to the protein involving the mannosyl residue are listed. P = primary and S = secondary binding sites in WGA.

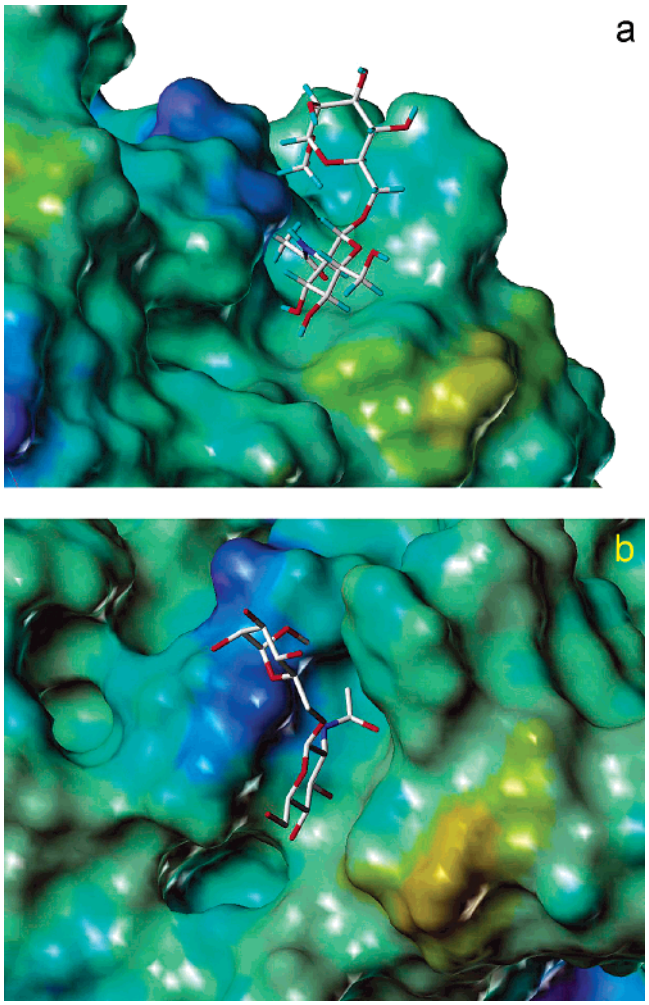

FIGURE 6: Molecular model (GT_C1) of **1** in a primary binding site of WGA. The protein is color-coded according to the electrostatic potential from blue to red for negative to positive charges, respectively. The conformation of ω in **1** is *gt*, and ψ is around −100°. (a) Side view and (b) top view with omission of the hydrogen atoms. Note the spatial proximity between the two methyl groups in the bound conformation of the disaccharide.

are preferred, namely, the *gt* and *gg* orientation of ω, with ψ in the extended antiperiplanar orientation. Interestingly, this preferred orientation of the ψ torsion angle is not allowed in the WGA-binding site because of steric conflict. As a result, the lectin binds the disaccharide in an altered conformation.

Such a selection of secondary minima in the binding of oligosaccharides by lectins have been previously illustrated by X-ray crystallography (see review in ref *46*). In general, the lectins bind the most populated conformation in solution, but several counter examples have been observed. In solution, the β-D-Glc*p*NAc-(1→2)-α-D-Man*p* disaccharide adopts one main low-energy minimum that could not fit in the binding site of mannose-specific legume lectins. When part of a pentasaccharide bound to ConA, this linkage adopts a conformation distorted by about 50° in ψ, still on the edge of the low-energy region (*47*), while when part of an octasaccharide bound to *Lathyrus ochrus* lectin, a secondary minimum is adopted with an anti-ϕ conformation rotated away by 120° from the main minimum (*48*). Similarly, the glycosidic linkage of the β-D-Gal*p*-(1→4)-β-D-Glc*p* fragment in GM_1 displays two distinct rotamers about the ϕ torsion angle when bound to the cholera toxin (*49*), and the α-NeuAc-(2→3)-β-D-Gal*p* linkage adopts a conformation different from the one observed in other sialic-acid-binding proteins. Interestingly, a recent NMR study that used the same approach as the one described here, demonstrated differential conformer selection of the α-NeuAc-(2→3)-β-D-Gal*p* fragment in the complexes between either sialyllactose or GM_1 and galectin-1 (*50*).

Future interaction studies on protein−carbohydrate complexes should further clarify the importance of different contributions such as hydrogen bonding, hydrophobic stacking, and the conformational flexibility of the ligand to specificity and selectivity of binding and subsequent signaling processes.

ACKNOWLEDGMENT

We thank Dr. C. S. Wright for the WGA coordinates.

REFERENCES

1. Dennis, J. W., Laferte, S., Yagel, S., and Breitman, M. L. (1989) Asparagine-linked oligosaccharides associated with metastatic Cancer, *Cancer Cells 1*, 87−92.
2. Dennis, J. W., Granovsky, M., and Warren, C. E. (1999) Glycoprotein glycosylation and cancer progression, *Biochim. Biophys. Acta 1473*, 21−34.

3. Dennis, J. W., Laferté, S., Waghorne, C., Breitman, M. L., and Kerbel, R. S. (1987) β1−6 Branching of Asn-linked oligosaccharides is directly associated with metastasis, *Science 236*, 582−585.
4. Hebert, E., and Monsigny, M. (1993) Oncogenes and expression of endogenous lectins and glycoconjugates, *Biol. Cell 79*, 97−109.
5. Fernandes, B., Sagman, U., Auger, M., Demetrio, M., and Dennis, J. W. (1991) β1−6 Branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia, *Cancer Res. 51*, 718−723.
6. Wojciechowicz, D. C., Park, P. Y., and Paty, P. B. (1995) β1−6 Branching of N-linked carbohydrate is associated with *K-ras* mutation in human colon carcinoma cell lines, *Biochem. Biophys. Res. Commun. 212*, 758−766.
7. Yamamoto, H., Swoger, J., Greene, S., Saito, T., Hurh, J., Sweeley, C., Leestma, J., Mkrdichian, E., Cerullo, L., Nishikawa, A., Ihara, Y., Taniguchi, N., and Moskal, J. R. (2000) β1,6-*N*-acetylglucosamine-bearing *N*-glycans in human gliomas: Implications for a role in regulating invasivity, *Cancer Res. 60*, 134−142.
8. Cummings, R. D., Trowbridge, I. S., and Kornfeld, S. (1982) A mouse lymphoma cell line resistant to the leukoagglutinating lectin from *Phaseolus vulgaris* is deficient in UDP-GlcNAc:α-D-mannoside β1,6-*N*-acetylglucosaminyltransferase, *J. Biol. Chem. 257*, 13421−13427.
9. Aub, J. C., Tieslau, C., and Lankester, A. (1963) Reactions of normal and tumor cell surfaces to enzymes, I. Wheat-germ lipase and associated mucopolysaccharides, *Proc. Natl. Acad. Sci. U.S.A. 50*, 613−619.
10. Wright, C. S., and Raikhel, N. (1989) Sequence variability in three wheat germ agglutinin isolectins: Products of multiple genes in polyploid wheat, *J. Mol. Evol. 28*, 327−336.
11. Wright, C. S. (1980) Crystallographic elucidation of the saccharide binding mode in wheat germ agglutinin and its biological significance, *J. Mol. Biol. 141*, 267−291.
12. Wright, C. S. (1984) Structural comparison of the two distinct sugar binding sites in wheat germ agglutinin isolectin II, *J. Mol. Biol. 178*, 91−104.
13. Wright, C. S. (1990) 2.2 Å resolution structure analysis of two refined *N*-acetylneuraminyl-lactose-wheat germ agglutinin isolectin complexes, *J. Mol. Biol. 215*, 635−651.
14. Wright, C. S. (1992) Crystal structure of a wheat germ agglutinin/glycophorin-sialoglycopeptide receptor complex. Structural basis for cooperative lectin-cell binding, *J. Biol. Chem. 267*, 14345−14352.
15. Muraki, M., Ishimura, M., and Harata, K. (2002) Interactions of wheat-germ agglutinin with GlcNAcβ1,6Gal sequence, *Biochim. Biophys. Acta 1569*, 10−20.
16. Bains, G., Lee, R. T., Lee, Y. C., and Freire, E. (1992) Microcalorimetric study of wheat germ agglutinin binding to *N*-acetylglucosamine and its oligomers, *Biochemistry 31*, 12624−12628.
17. Midoux, P., Grivet, J.-P., and Monsigny, M. (1980) Lectin-sugar interactions: The binding of 1-*O*-methyl-di-*N*-trifluoroacetyl-β-chitobioside to wheat germ agglutinin, *FEBS Lett. 120*, 29−32.
18. Umemoto, K., Oikawa, S., Aida, M., and Sugawara, Y. (1988) Intermolecular nuclear Overhauser effect and atomic pair potential approaches to wheat germ agglutinin-sugar binding, *J. Biomol. Struct. Dyn. 6*, 593−608.
19. Nagahora, H., Harata, K., Muraki, M., and Jigami, Y. (1995) Site-directed mutagenesis and sugar-binding properties of the wheat germ agglutinin mutants Tyr73Phe and Phe116Tyr, *Eur. J. Biochem. 233*, 27−34.
20. Kawashima, H., Sueyoshi, S., Li, H., Yamamoto, K., and Osawa, T. (1990) Carbohydrate binding specificities of several poly-*N*-acetyllactosamine-binding lectins, *Glycoconjugate J. 7*, 323−334.
21. Fukuda, M. (1991) Leukosialin, a major O-glycan-containing sialoglycoprotein defining leukocyte differentiation and malignancy, *Glycobiology 1*, 347−356.
22. Kjellberg, A., and Widmalm, G. (1999) A conformational study of the vicinally branched trisaccharide β-D-Glc*p*-(1→2)[β-D-Glc*p*-(1→3)]α-D-Man*p*-OMe by nuclear Overhauser effect spectroscopy (NOESY) and transverse rotating-frame Overhauser effect spectroscopy (TROESY) experiments: Comparison to Monte Carlo and Langevin dynamics simulations, *Biopolymers 50*, 391−399.
23. Kupce, E., Boyd, J., and Campbell, I. D. (1995) Short selective pulses for biochemical applications, *J. Magn. Reson., Ser. B 106*, 300−303.
24. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) Rapid recording of 2D NMR spectra without phase cycling: Application to the study of hydrogen exchange in proteins, *J. Magn. Reson. 85*, 393−399.
25. Stott, K., Keeler, J., Van, Q. N., and Shaka, A. J. (1997) One-dimensional NOE experiments using pulsed field gradients, *J. Magn. Reson. 125*, 302−324.
26. Imberty, A., Bettler, E., Karababa, M., Mazeau, K., Petrova, P., and Pérez, S. (1999) Building sugars: The sweet part of structural biology, in *Perspectives in Structural Biology* (Vijayan, M., Yathindra, N., and Kolaskar, A. S., Eds.) pp 392−409, Indian Academy of Sciences and Universities Press, Hyderabad, India.
27. Clark, M., Cramer, R. D., III, and van Opdenbosch, N. (1989) Validation of the general purpose Tripos 5.2 force field, *J. Comput. Chem. 10*, 982−1012.
28. Wright, C. S., and Kellogg, G. E. (1996) Differences in hydropathic properties of ligand binding at four independent sites in wheat germ agglutinin−oligosaccharide crystal complexes, *Protein Sci. 5*, 1466−1476.
29. Marchessault, R. H., and Pérez, S. (1979) Conformations of the hydroxymethyl group in crystalline aldohexopyranoses, *Biopolymers 18*, 2369−2374.
30. Imberty, A., Delage, M. M., Bourne, Y., Cambillau, C., and Pérez, S. (1991) Data bank of three-dimensional structures of disaccharides: Part II. *N*-acetyllactosaminic type *N*-glycans. Comparison with the crystal structure of a biantennary octasaccharide, *Glycoconjugate J. 8*, 456−483.
31. Pérez, S. (1978) Analyse crystallographique de structures polymères: conception et critique de nouveaux systèmes d'information, Ph.D. Thesis, Université Scientifique et Médicale, Grenoble, France.
32. Jansson, P.-E., Kenne, L., and Kolare, I. (1994) NMR studies of some (1 → 6)-linked disaccharide methyl glycosides, *Carbohydr. Res. 257*, 163−174.
33. Laatikainen, R., Niemitz, M., Weber, U., Sundelin, J., Hassinen, T., and Väpsäläinen, J. (1996) General strategies for total-line shape-type spectral analysis of NMR spectra using integral-transform iterator, *J. Magn. Reson., Ser. A 120*, 1−10.
34. Laatikainen, R., Niemitz, M., Malaisse, W. J., Biesemans, M., and Willem, R. (1996) A computational strategy for the deconvolution of NMR spectra with multiplet structures and constraints: Analysis of overlapping ^{13}C-^{2}H multiplets of ^{13}C enriched metabolites from cell suspensions incubated in deuterated media, *Magn. Reson. Med. 36*, 359−365.
35. Stenutz, R., Carmichael, I., Widmalm, G., and Serianni, A. S. (2002) Hydroxymethyl group conformation in saccharides: Structural dependencies of $^{2}J_{HH}$, $^{3}J_{HH}$, and $^{1}J_{CH}$ spin−spin coupling constants, *J. Org. Chem. 67*, 949−958.
36. Keepers, J. W., and James, T. L. (1984) A theoretical study of distance determinations from NMR. Two-dimensional nuclear Overhauser effect spectra, *J. Magn. Reson. 57*, 404−426.
37. Kronis, K. A., and Carver, J. P. (1982) Specificity of isolectins of wheat germ agglutinin for sialyloligosaccharides: A 360-MHz proton nuclear magnetic resonance binding study, *Biochemistry 21*, 3050−3057.
38. Espinosa, J. F., Asensio, J. L., García, J. L., Laynez, J., Bruix, M., Wright, C., Siebert, H.-C., Gabius, H.-J., Cañada, F. J., and Jiménez-Barbero, J. (2000) NMR investigations of protein-carbohydrate interactions. Binding studies and refined three-dimensional solution structure of the complex between the B domain of wheat germ agglutinin and *N*,*N*′,*N*″-triacetylchitotriose, *Eur. J. Biochem. 267*, 3965−3978.
39. Arepalli, S. R., Glaudemans, C. P. J., Daves, G. D., Jr., Kovac, P., and Bax, A. (1995) Identification of protein-mediated indirect NOE effects in a disaccharide-Fab′ complex by transferred ROESY, *J. Magn. Reson., Ser. B 106*, 195−198.
40. Weimar, T., and Peters, T. (1994) *Aleuria aurantia* agglutinin recognizes multiple conformations of α-L-Fuc-(1→6)-β-D-GlcNAc−OMe, *Angew. Chem. Int. Ed. Engl. 33*, 88−91.
41. Almond, A., and Duus, J. Ø. (2001) Quantitative conformational analysis of the core region of *N*-glycans using residual dipolar couplings, aqueous molecular dynamics, and steric alignment, *J. Biomol. NMR 20*, 351−363.

42. Forsén, S., and Hoffman, R. A. (1963) Study of moderately rapid chemical exchange reactions by means of nuclear magnetic double resonance, *J. Chem. Phys. 39*, 2892−2901.
43. Mayer, M., and Meyer, B. (2001) Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor, *J. Am. Chem. Soc. 123*, 6108−6117.
44. Meyer, B., and Peters, T. (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors, *Angew. Chem. Int. Ed. 42*, 864−890.
45. Rinnbauer, M., Ernst, B., Wagner, B., Magnani, J., Benie, A. J., and Peters, T. (2003) Epitope mapping of sialyl Lewis^x bound to E-selectin using saturation transfer difference NMR experiments, *Glycobiology 13*, 435−443.
46. Imberty, A., and Pérez, S. (2000) Structure, conformation, and dynamics of bioactive oligosaccharides: Theoretical approaches and experimental validations, *Chem. Rev. 100*, 4567−4588.
47. Moothoo, D. N., and Naismith, J. H. (1998) Concanavalin A distorts the β-GlcNAc-(1→2)-Man linkage of β-GlcNAc-(1→2)-α-Man-(1→3)-[β-GlcNAc-(1→2)-α-Man-(1→6)]-Man upon binding, *Glycobiology 8*, 173−181.
48. Bourne, Y., Rougé, P., and Cambillau, C. (1992) X-ray structure of a biantennary octasaccharide−lectin complex refined at 2.3 Å resolution, *J. Biol. Chem. 267*, 197−203.
49. Merritt, E. A., Sarfaty, S., Jobling, M. G., Chang, T., Holmes, R. K., Hirst, T. R., and Hol, W. G. J. (1997) Structural studies of receptor binding by cholera toxin mutants, *Protein Sci. 6*, 1516−1528.
50. Siebert, H. C., André, S., Lu, S. Y., Frank, M., Kaltner, H., van Kuik, J. A., Korchagina, E. Y., Bovin, N., Tajkhorshid, E., Kaptein, R., Vliegenthart, J. F. G., von der Lieth, C.-W., Jiménez-Barbero, J., Kopitz, J., and Gabius, H.-J. (2003) Unique conformer selection of human growth-regulatory lectin galectin-1 for ganglioside GM_1 versus bacterial toxins, *Biochemistry 42*, 14762−14773.

BI0499011